Comparison of autopsy, clinical and death certificate diagnosis with particular reference to lung cancer

A review of the published evidence

Author : P.N. Lee

Date : November 1992

1. INTRODUCTION

This document gives a brief summary of a large number of papers relating to accuracy of diagnosis. I have restricted attention to papers which concern lung cancer, all cancer, or all deaths, and have not considered papers addressing accuracy of diagnosis of other specific diseases. I have also restricted attention to papers in the English language. For each paper of interest I have attempted to summarize relevant information in a paragraph or two and to organize those paragraphs into sections under common headings ordered chronologically. With the exception of a summary table for section 2.8, I have not so far attempted to interpret or summarize. The main intention was to familiarize myself (and others) with relevant evidence, to assist in writing up papers describing results from the Budapest study currently being analyzed. The possibility is being considered of preparing a paper from the material described here for publication in a journal.

Any comments on the material presented, or references of any missing relevant papers I may have overlooked, are welcome, as are ideas for the most useful form a review might take, and where it might be submitted.

2028470266

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

2. AUTOPSY

2.1 The Autopsy

According to the Oxford English Dictionary the word "autopsy" is defined as follows: "Seeing with one's own eyes, personal observation. Dissection of a dead body, so as to ascertain by actual inspection esp. the cause or seat of disease; post-mortem examination".

McPhee and Bottles (1985) summarize the procedure as follows: "After examination for exterior abnormalities, including growths, wounds, and surgical scars, a surgical incision is made to allow gross inspection of the internal organs. The brain and spinal cord are then removed. The removed organs are weighed, measured, and examined, and incisions are made in each to identify fine details. Next, tissue samples are obtained for light and electron microscopy, cultures, and toxicologic analysis. After autopsy, the skin incisions are sewn up and the body released to the mortuary for burial. Actual dissection takes two to four hours depending on whether the autopsy is total or limited, but the final report is often not finished for months. About three quarters of the autopsies performed in the United States are total."

McPhee and Bottles also summarize the history of the autopsy, the roots of which "can be traced back some 3,500 years to the Babylonians who attempted to foretell the future by examining the liver or entrails of sacrificed animals". "The general attitude of early Church leaders toward autopsy was unfavourable", but "when Pope Alexander died suddenly in 1410, an autopsy was performed", and

2028470267

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

"Pope Sixtus IV issued a bill permitting autopsy by students at Padua and Bologna". "In 1679, Bonetus published a collection of more than 3,000 autopsy reports by some 450 physicians", and "in the 18th century Morgagni began to correlate clinical observations with autopsy findings". "In the late 19th century, Rokitansky is alleged to have performed 30,000 autopsies. Only at the end of his career, however, was the microscope available. Virchow, 17 years younger, grew up with the microscope, and published a book on autopsy technique". "Early in the 20th century, many outstanding clinicians spent considerable time in the autopsy room" and "were frequently good pathologists".

However, "the autopsy rate in the US has fallen dramatically from about 50 percent of deaths after World War II to a mere 15 percent today". McPhee and Bottles cited seven reasons for this decline, which can be summarized as follows:

(i) pathologists have many other responsibilities,

(ii) autopsy is often relegated to the most junior members of pathology departments,

(iii) clinicians are requesting fewer autopsies, some for fear of malpractice suits,

(iv) the process of obtaining consent for autopsy is cumbersome,

(v) family members are in emotional turmoil and do not perceive its value,

(vi) hospital accreditation no longer requires a 20% autopsy rate, and

(vii) autopsies are expensive and are generally not reimbursable.

02084702G8

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

It should also be noted that in the United Kingdom the National Health Service does not provide any money for autopsy.

Prutting (1967) referred to Mencken's definition of a judge as "a law student who grades his own papers", and states that "a similar view might be taken of the physician who fails to submit his diagnostic skills to that impartial grader, the autopsy". He notes that "the case for the autopsy rests solidly on the well-documented fact that antemortem diagnosis is frequently refuted, clarified, modified, or elaborated by postmortem examination. The case against the autopsy rests on shakier grounds, principally lay resistance and the lack of time by professional personnel to perform the autopsy".

2.2 Autopsy rates

The WHO in their World Health Statistics Annual routinely provide "background information on mortality statistics" which includes information on the percentage of deaths for which an autopsy was performed. Table 1 summarizes latest available data in 1991 for a number of countries on the autopsy rate overall and by age (where available). Of the 27 countries, the highest rate is for Hungary (49%), the lowest for Japan (4%), with the median 20%. For all countries where data are available, the lowest rates are in the oldest age group.

Kennaway and Kennaway (1936) examined death certificates for all 8808 cases of lung cancer in males from England and Wales for the years 1921-32. A post mortem had been carried out for 25%, the rate rising from 12% in 1921-23 to 26% in 1930-32. The authors comment on the "high proportion of cases thought to have died of

2028470269

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

TABLE 1

Data cited by WHO on autopsy rates in 1991

Country	% of deaths for which an autopsy was performed			
	Total	By age		
		15-44	45-64	65+
Hungary	49			
Iceland	38			
Sweden	37	76	56	36
Finland	36	82	52	27
Austria	34	47	46	31
Denmark	32	30*	27*	17*
USSR	31			
Czechoslovakia	29			
England & Wales	26	62	36	22
New Zealand	26	71	35	18
Bulgaria	25			
East Germany	25	48	34	14
Australia	21	70	29	12
Canada	20	56	29	13
Switzerland	19	35	27	16
Singapore	16			
Scotland	15	57	24	11
Kuwait	15	48	13	3
Norway	14			
USA	12	15	16	5
Northern Ireland	11	56	17	5
Poland	9			
West Germany	8			
Netherlands	8	14	11	7
Eire	7	13	13	4
Malta	6	39	9	2
Japan	4	8	7	3

(* Age groups 35-54, 55-70, 75+)

cancer of the lung ... now subjected to autopsy". For larynx cancer rates were only 6% over the whole period rising to 8% in 1930-32.

Beadenkopf et al (1965), for deaths of 15+ year olds in Albany County, New York during 1955-1959, compared various demographic characteristics of all 12,930 decedents, those 2574 who had died at Albany Medical Center Hospital and those 1586 who had been autopsied. The likelihood that a person who died in the county had

02847027O

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

been hospitalized and autopsied decreased markedly with age, particularly in females. After adjustment for age and sex, those autopsied were considerably less likely to be certified as having died of heart disease than other diseases, and to be Jews rather than Protestants. Negros, the divorced, residents of small towns or villages, and native-born individuals were more likely to be autopsied than those in other categories of race, marital status, place of residence and place of birth.

Kunitz and Edland (1973) carried out a study of the autopsy rate in Monroe County, New York, based on deaths occurring in the years 1960-1970. The rate was shown to be higher in younger people, in males, and in blacks and for deaths occurring at the University Hospital.

Guptill (1976) carried out a study of the social factors associated with permission for autopsies to be conducted in Michigan in 1965-66, interviewing the next of kin of 475 county residents aged 40 or over who died in one of the 3 major hospitals in Washtenaw County. The older the patient, the lower the probability permission for an autopsy would be requested and the lower the probability permission would be granted if requested. Autopsies were only conducted in 45% of cases where the patient was 75+ as against 76% where the patient was 65+. Permission to conduct an autopsy was more likely to occur if the next-of-kin was younger and a closer relative.

Cameron (1981) deplored the low rate of autopsy in Scotland and suggested that it should be possible to arrange that autopsies be carried out on a sample - say 20% - of deaths in addition to those

2028470271

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

which would be asked for in any case. He noted that "some may contend that difficulty in obtaining permission will block any attempt to increase autopsy numbers. The evidence suggests otherwise: some clinical units achieve a 90% autopsy rate, while others in the same geographic area are apparently content with 5% to 10%." Cameron stated that a greater use of the autopsy would "have no small effect on the accuracy of death certification and hospital statistics. Both are wildly inaccurate in Scotland."

Based on 99,145 death certificates from New Jersey in 1979 and 1980, Ahronheim et al (1983) demonstrated a striking age difference in the frequency with which autopsies were carried out. The peak (82.5%) occurred in the third decade of life, the nadir (2.4%) by age 90 years. The trend was similar in both sexes, though at all ages men were more likely to undergo autopsies than women. The age-related decline was evident for all causes except those for which an autopsy was mandatory by law (homicide, suicide, transport accidents).

McFarlane (1987) compared the demographic features of 4122 adult patients who died at the University of Kansas Medical Center and underwent autopsy with the remaining 3271 patients who died and did not undergo autopsy. He reported no "clinically relevant" differences between autopsied and non-autopsied patients as regards age, sex, or race, although the large sample size made differences statistically significant. For example, the mean age of those undergoing necropsy was only 2 years younger, in each sex, than that of those not undergoing necropsy. The author concluded that his

2028470272

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

findings suggest that "a demographic bias does not distort necropsy population data compared with data of the population of hospital records".

Declich et al (1991) studied factors associated with an autopsy being conducted in an area of Italy, the Unita Socio-Sanitaria Locale, where, over the period 1985-88, 3041 deaths occurred of which 321 (10.6%) were autopsied. Factors associated with autopsy were death at a young age, short length of hospitalization, and death from a cardiovascular disease, but not sex. Deaths from neoplasms (6.1%) and from lung cancer (5.7%) were autopsied less often than average.

A report by a Joint Working Party of the Royal Colleges of Pathologists, Physicians and Surgeons (1991) notes a progressive decline in autopsy rates throughout the world. It notes that "though religious objection is often cited as a cause ... few religions take a stand against autopsies". It also points out that though "it is often thought that families may suffer extra distress when an autopsy is reported", a study by McPhee et al (1986) of the families of patients undergoing an autopsy had actually shown that 88% considered they had benefited. Although there have been advances in diagnostic techniques they note that discrepancies between clinical and autopsy diagnosis have remained around the 10% level. Costs are also an issue referred to. They cite a study by Nemetz et al (1987) noting that "the decline in the autopsy rate in the Mayo Clinic has been associated with the failure of the autopsy service to provide income to the institution, even though it has been shown to play an advantageous role in quality control, cost containment

2028470273

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

and efficient allocation of resources". They recommend that "the autopsy service in hospitals must be adequately funded with a requirement to maintain fundings at a level to sustain effective clinical audit", and make a number of other recommendations concerning the performance and use of autopsies. These recommendations include sampling of general hospital deaths in which there is no perceived prior necessity to perform an autopsy at the rate of "at least 10%".

2.3 Accuracy of autopsy diagnosis

Wilson (1966) has pointed out that "in justice and honesty it must be admitted that in a certain number of patients - about 20% - more or less mystery and confusion still prevailed after the post-mortem examinations".

Jimenez et al (1975) reexamined all the available slides in 465 male autopsy cases previously certified as bronchogenic carcinoma, the deaths occurring at the Brooklyn Veterans Hospital over the period 1964-73. In 152 cases the diagnosis was only established by partial autopsy or by needle biopsy, considered insufficient to establish the lung as the primary site of origin. In 253 of the remaining 313, which had sufficient histologic material to permit complete microscopic examination, the diagnosis was confirmed. Among the 60 cases erroneously diagnosed, there were 14 cases in which no evidence of malignancy was found at autopsy (the past history revealing there had been a previous resection), and 46 cases in which pulmonary metastases from other primary sites simulated bronchogenic carcinoma.

202847027A

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

Cameron (1991) noted that when he and his colleages completed their work in Edinburgh some years ago, revealing "dramatic differences between clinical diagnoses and findings at autopsy", there was a "deafening silence", with "virtually no reaction from either clinicians or pathologists". He cited, with evident horror, the reaction of a clinician when told of the "lamentable fall" in the autopsy rate, which was "I am sure this merely reflects the increase in diagnostic techniques that allow a definitive ante-mortem diagnosis to be accurately established, thus happily eliminating the need of necropsy" (Cameron's emphasis). He summarized data from a number of studies indicating considerable inaccuracy of clinical diagnosis, and criticized as incomprehensible a recent authoritative booklet called "Medical Audit" (Royal College of Physicians, 1989), because it "nowhere" suggested that "clinical diagnoses may be suspect or challengeable" and "nowhere" indicated that "the autopsy has any contribution to make. Cameron considered that "it is difficult to know how to penetrate this carapace of neglect and indifference".

2.4 Measurement of agreement between clinical and necropsy diagnosis

Saracci (1991a,1991b) considers the 2 x 2 table for presence or absence of a disease at post mortem and clinically. He presents arguments why the sensitivity and specificity are indices that better monitor clinical diagnosis performance than are the agreement rate and the confirmation rate. He also points out that there is a need to have data concerning the likely size of errors in postmortem diagnosis. Specificity and sensitivity of clinical diagnosis

2028470275

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

compared to true diagnosis will not be accurately measured by comparing clinical diagnosis to post mortem diagnosis unless errors in post mortem diagnosis are very small indeed.

2.5 Classifying discrepancy between clinical and necropsy diagnosis

Anderson (1984) described a method for classifying a discrepancy between the clinical and necropsy diagnosis:

Class 1 was a discrepant major diagnosis with adverse impact on survival,

Class 2 was a discrepant major diagnosis with equivocal impact on survival,

Class 3 was a discrepant minor diagnosable disease that was not directly related to cause of death but either was asymptomatic and should have been treated or would have eventually affected prognosis, and

Class 4 was a discrepant nondiagnosable minor disease of possible genetic or epidemiological importance.

He also presented a table of 12 published autopsy studies showing percentages showing substantial discrepancies (including only those with substantial impact on patient survival) varying between 4% and 18.7% (median 9.6%).

2.6 National autopsy data available on computer

In 1959 the Japanese Pathological Society established a national autopsy registration system, data from which are published every year as an annual. Urano et al (1982) point out that computerization of the autopsy data started in 1974, and describe the format by which the data are recorded. They also give detailed tables showing the age distribution of various diseases, the joint

2028470276

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

distribution of tumours by site and histological type, and the combination of primary sites in double malignancies. The autopsy rate was noted to be only 3.9% over the period 1974-79 where data had been computerized.

2.7 Reasons for clinical misdiagnosis

Onuigbo (1971) provides a useful summary paper emphasizing the difficulties of diagnosis of lung cancer in the 19th century (X-rays were not discovered by Röntgen until 1895). 51 papers dated before 1900 are cited, including one in 1793!

Rosenblatt et al (1966) studied 380 cases of extrathoracic carcinoma necropsied at the Doctor's Hospital, New York between 1945 and 1964. 189 (44.7%) had pulmonary metastases, the most frequent being colon (39), breast (41) and pancreas (28). 97 (25.5%) had bronchial involvement, approximately 50% of the cases with metastases. 39 (10.3%) presented features compatible with the diagnosis of bronchogenic carcinoma, approximately 40% of the cases with bronchial involvement. 14 of these 39 cases were metastases of pancreatic carcinomas, kidney carcinomas being next most frequent (6 cases). The authors conclude that "the simulation of bronchogenic carcinoma by metastatic tumours occurs with sufficient frequency to challenge the diagnostic accuracy of death certificates as bronchogenic carcinoma without necropsy exclusion of other primary sites".

Willis (1967) describes differing methods of diagnosing cancer with increasing likelihood of precision:

Grade 1. Diagnosis by simple physical examination only.

2028470277

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

Grade 2. Diagnosis by physical examination supplemented by special methods.

Grade 3. Diagnosis by histological examination of surgically removed tissues.

Grade 4. Diagnosis by necropsy, (a) without histological confirmation, (b) with histological confirmation.

For Grade 1 he notes the difficulty of diagnosing common internal tumours, and states that "the supposedly increased frequency of such tumours during recent years is undoubtedly due largely to increasing recognition by practitioners of the need for enlisting the aid of special methods of investigation in suspect areas", and that "the nature of the community greatly affects the accuracy of diagnosis". Although methods under Grade 2, such as X-ray, "greatly enhance the accuracy of diagnosis of many internal tumours" ... "in the absence of pathological examination of excised tumours of or biopsy fragments, an important degree of diagnostic tumours still obtains". Histological examination of surgically removed tissues (Grade 3) may sometimes provide a "final and complete" diagnosis, but may often leave the site of the primary tumour undetermined. Any post-mortem examination without histological confirmation is considered incomplete by Willis. "Not infrequently, microscopic examination surprises the clinician and pathologist by showing the falsity of the initial post-mortem diagnosis, even in cases where this appeared certain". He warns, however, that "not all histopathologists are of equal ability and experience, questionable diagnoses ... are far from rare".

20284278

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

Feinstein and Wells (1974) carried out a study based on 654 patients at Yale-New Haven and West Haven VA Hospitals during the 12 year period 1953-1964 with a post-mortem diagnosis of lung cancer, of which 102 had not been diagnosed during life. The rate of non-diagnosis during life was found to be lower with increasing amount of smoking, ranging from 38% in non-smokers to 10% in smokers of more than 1 pack per day. The authors concluded that "the data suggested that the more patients smoked, the more likely they were to have their lung cancer detected during life". Based on those patients and also a further 677 patients not autopsied, they also reported that the "search rate" in ordering sputum Papanicolau ("Pap") smears was higher the more the patient smoked, ranging from 58% in non-smokers to 84% in smokers of more than 1 pack per day. The search rate was also noted to be higher in those with a cough, in men, and in more recent patients.

Alderson et al (1983) in an editorial on death certification in the British Medical Journal notes that "many reports in Britain and overseas have shown that after necropsy the pathologist's opinion about the cause of death differs from that of the clinician in charge of the patient in a substantial proportion of cases. This divergence is much greater in the elderly, in whom multiple diseases are common, and especially when neoplasms or cardiovascular or respiratory diseases were thought in life to have been responsible for the last illness." They also point out that "death may, however, have resulted from bronchial disturbances or from cardiac arrest or an arrhythmia which cannot be recognized at necropsy, and conversely structural disease evident to the pathologist may have no

2028470279

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

relevance to the fatal outcome. The less consultation there is between clinician and pathologist the wider will be their disagreement."

Wells and Feinstein (1988) carried out a case-control study of men or women aged 25 or older attending Yale-New Haven Hospital. The 385 cases had received a sputum Pap smear test during the period 1977 to 1980. Controls were individually matched on age, admission diagnosis and date of admission but did not receive a Sputum Pap Smear test. Patients were not included as cases or controls if they had characteristics which rendered a sputum Pap smear test obligatory, unnecessary, or unproductive. The odds ratio for having a Pap smear test was 1.90 in men compared to women, 2.78 in those with a recent cough compared to those without one, 3.71 in those with a chronic cough compared to those without one, and 2.86 in those who had ever smoked compared to those who had never done so. All these differences were extremely highly statistically significant and largely independent of each other. The authors "suggest that women and non-smokers may be deprived of appropriate diagnosis and therapy" unless diagnosis of lung cancer is "guided mainly by radiographic findings and presenting manifestations" and note that "detection bias has probably led to an excessively elevated magnitude for the cigarette smoking-lung cancer association and to a falsely low estimate of incidence rates in women".

In an abstract, Wells and Feinstein (1992) describe results from a cohort of 1266 patients with primary lung cancer. They noted that median age showed an inverse monotonic gradient with smoking, ranging from 69 for nonsmokers down to 59 for smokers of more than

2028470280

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

two packs of cigarettes per day. In the localized lesions of TNM stage I this negative gradient persisted if suspicious clinical symptoms were absent, but vanished if they were present. Regardless of symptoms, the gradient was strengthened in metastatic TNM stages "where diverse sources might be considered for a primary, but where smoking might particularly suggest the lung". The authors conclude from this cohort, and also from two other smaller cohorts of 132 and 178 patients, which showed similar results, that "detection bias plays a prominent, although often neglected, role in the work-up decisions that precede and lead to a diagnosis of lung cancer".

2.8 Differences between clinical and necropsy diagnoses - studies not specifically mentioning lung cancer

JAMA (1914) refers to a report based on a comparison of clinical and necropsy diagnoses made in New York by Oertel. Among 388 cases, the clinical diagnoses were confirmed in 22.4% (87), the diagnoses were correct but the necropsies disclosed additional important lesions in 29.9% (116), the diagnoses were partially correct but the necropsies revealed other important lesions in 13.9% (107), the clinical diagnoses were not confirmed in 27.6% (107), and in 6.2% (24) no clinical diagnosis was recorded. The report concluded that "too great reliance is placed on inexperienced house physicians and interns, and also that the current knowledge necessary to make clinical diagnoses which shall approach accuracy is insufficient".

2028470281

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000

Karsner et al (1919) compared clinical and autopsy diagnosis of 600 cases in two large hospitals in Cleveland, Ohio. In 50 (8%) of the cases there were what was considered gross errors in the clinical diagnosis. The overall error rates, including also minor errors, were presented only by organ or part of the body. For lungs and pleura the chance of a correct diagnosis was noted to be 38%.

Cleland (1942) compared clinical and autopsy diagnosis for 2500 deaths occurring at the Royal Adelaide Hospital over about a 20 year period. Of these 67.8% (1697) were considered "correctly diagnosed before death or at any rate reasonably so", 12.4% (310) were considered partly correct, 14.6% (362) were considered wrong and in 5.2% (131) no diagnosis was made, or the cause of death still remained obscure after post-mortem examination. Cancer was correctly diagnosed clinically in 166 cases, though the site was wrongly inferred in 10. It was present but not diagnosed clinically in 27 cases, and was diagnosed or suspected but not present in 22.

Gruver and Freis (1957) studied the clinical records of 1106 patients autopsied over the period 1947-1953 at the Veterans Administration Hospital in Washington, DC, where the autopsy rate was 7.2%. After excluding cases where "a precise diagnosis was impossible, e.g. metastatic carcinoma in which the primary site was incorrect" they reported that 64 (6%) was misdiagnosed. They noted that misdiagnoses were more common in June and July when the resident staff is changed and the students and other personnel are on vacation. 15 of the 64 misdiagnoses were neoplasms, of which 2 were of the lung.

2028470282

Source: https://www.industrydocuments.ucsf.edu/docs/xscl0000